Appendix A: Technical Notes

Age-Adjustment

Age-adjusted incidence rates were developed using the direct method. They were standardized to the age distribution of the United States 1970 and 1940 populations. Following the age-adjustment procedures used by the National Cancer Institute, which uses the US 1970 standard population for age-adjustment, we used five year age groups in calculating age-adjusted rates with the 1970 US standard population. For age-adjustment with the 1940 US standard population, we followed the methods of the National Center for Health Statistics which uses 10 year age groups from age 5 through 85. The age distributions of the US standard populations are shown below.

US Standard Population Proportions

1970 age group	1970 proportion	1940 age group	1940 proportion
0 - 4	0.0844	<1	0.0160
5 - 9	0.0982	1 - 4	0.0641
10 - 14	0.1023	5 - 14	0.1703
15 - 19	0.0938	15 - 24	0.1817
20 - 24	0.0806	25 - 34	0.1621
25 - 29	0.0663	35 - 44	0.1392
30 - 34	0.0562	45 - 54	0.1178
35 - 39	0.0547	55 - 64	0.0803
40 - 44	0.0590	65 - 74	0.0484
45 - 49	0.0596	75 - 84	0.0173
50 - 54	0.0546	85+	0.0028
55 - 59	0.0491		
60 - 64	0.0424		
65 - 69	0.0344		
70 - 74	0.0268		
75 - 79	0.0189		
80 - 84	0.0112		
85+	0.0074		

Direct method of age adjustment

Multiply the age-specific rates in the target population by the age distribution of the standard population.

$$\hat{R} = \sum_{i=1}^{m} s_i (d_i / P_i) = \sum_{i=1}^{m} w_i d_i$$

Where m is the number of age groups, d_i is the number of deaths in age group i, P_i is the population in age group i, and s_i is the proportion of the standard population in age group i. This is a weighted sum of Poisson random variables, with the weights being (s_i / P_i).

Confidence Intervals

Confidence intervals for the age-adjusted rates were calculated with a method based on the gamma distribution (Fay and Feuer, 1997). This method produces valid confidence intervals even when the number of cases is very small. When the number of cases is large the confidence intervals produced with the gamma method are equivalent to those produced with the more traditional methods, as described by Chiang (1961) and Brillinger (1986). The formulas for computing the confidence intervals are given below. Although the derivation of this method is based on the gamma distribution, the relationship between the gamma and Chi-squared distributions allows the formulas to expressed in terms of quantiles of the Chi-squared distribution, which can be more convenient for computation.

$$\text{Lower Limit} = \frac{v}{2y}\left(c^2\right)^{-1}_{\frac{2y^2}{v}}\left(a/2\right)$$

$$\text{Upper Limit} = \frac{v + w_M^2}{2(y + w_M)}\left(c^2\right)^{-1}_{\frac{2(y+w_M)^2}{v+w_M^2}}\left(1 - a/2\right)$$

where y is the age-adjusted death rate, v is the variance as calculated as shown below, w_M is the maximum of the weights $s_i P_i$, $1 - a$ is the confidence level desired (e.g., for 95% confidence intervals, a = 0.05), and $\left(c^2\right)^{-1}_{x}$ is the inverse of the c^2 distribution with x degrees of freedom.

$$v = \sum_{i=1}^{m} d_i\left(s_i / P_i\right)^2$$

References

Brillinger, D. R. The natural variability of vital rates and associated statistics [with discussion]. *Biometrics* 42:693-734, 1986.

Chiang, C. L. Standard error of the age-adjusted death rate.*Vital Statistics, Special Reports* 47:271-285, USDHEW, 1961.

Fay, M.P. and E.J. Feuer. Confidence intervals for directly rates: a method based on the gamma distribution. *Stat Med*16:791-801, 1997